

Title	Developing control over the execution of scripts: The role of maintained hierarchical goal representations
Author(s)	Yanaoka, Kaichi; Saito, Satoru
Citation	Journal of Experimental Child Psychology (2017), 163: 87-106
Issue Date	2017-11
URL	http://hdl.handle.net/2433/235748
Right	© 2017. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/ ; The full-text file will be made open to the public on 01 November 2019 in accordance with publisher's 'Terms and Conditions for Self-Archiving'.; This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。
Type	Journal Article
Textversion	author

Journal of Experimental Child Psychology

Developing Control Over the Execution of Scripts:
The Role of Maintained Hierarchical Goal Representations

Kaichi Yanaoka and Satoru Saito

Address correspondence:

Kaichi Yanaoka

Graduate School of Education, Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Tel. and Fax: +81-75-753-3004

E-mail. yanaoka.kaichi.22a@st.kyoto-u.ac.jp

Yanaoka, K., & Saito, S. (2017). Developing control over the execution of scripts: The role of maintained hierarchical goal representations. *Journal of Experimental Child Psychology*, 163, 87–106.
<http://doi.org/10.1016/j.jecp.2017.06.008>

Abstract

The execution of a script often requires detecting and resolving conflict with a goal, particularly in non-routine situations. To take an example of taking a bus daily to work, if someone's usual bus delays and a bus for another destination comes first, the person must inhibit taking it and wait for the usual one. Young children can gradually acquire the ability to control the execution of scripts in such non-routine situations, but few studies have explored the control process involved. In two experiments, we investigated the role of developments in the maintenance of hierarchical goal representations and in executive functions. We measured the ability to control the execution of scripts using a task in which children helped a doll select items to wear (Yanaoka, 2014); clothing options were presented in an unexpected order in the non-routine situations. Four-year-olds could not flexibly control their execution of scripts in non-routine situations, although they could exogenously detect and resolve conflict if they were prompted to maintain a sub-goal. Five-year-olds endogenously controlled script execution based on a main goal, whereas sub-goal maintenance led them to rigidly control their performance. In addition, children's inhibition abilities were associated with their control of script execution. These findings indicate that the development of the control process underlying the execution of scripts in non-routine situations is partially dependent on the ability to maintain hierarchical goal representations.

Keywords: script, hierarchical goal representations, executive functions, preschool children, goal maintenance, order of actions

Introduction

In daily life, we are familiar with many scenarios in which we conduct action sequences to achieve short- and long-term goals. For example, individuals who go out to eat in a restaurant proceed through a stereotyped action sequence of events: they enter the restaurant, order food, eat, pay, and then leave. Similarly, in everyday life, we all acquire *scripts*; *i.e.*, event knowledge made up of action sequences (Nelson, 1986; Schank & Abelson, 1977). According to Schank and Abelson (1977), scripts are goal-directed and organized in a temporal-causal order. Moreover, scripts contain hierarchical structures in which both main goals and sub-goals are represented. Previous developmental studies (e.g., Fivush, 1984; Nelson, 1986) have demonstrated that 3- to 4-year-olds are capable of reporting well-organized scripts that play a fundamental role in their cognitive activities, such as imitation (e.g., Bekkering, Wohlschlager, & Gattis, 2000), pretend play (Furman & Walden, 1990), and text processing (Hudson & Slackman, 1990). In addition, it is notable that scripts guide goal-directed behavior over a relatively long period of time; even 3- to 4-year-olds may plan and enact natural action sequences based on their own scripts (Freier, Cooper, & Mareschal, 2015; Hudson & Fivush, 1991; Hudson, Shapiro, & Sosa, 1995; Loucks & Melzoff, 2013; Shapiro & Hudson, 2004; Yanaoka, 2014).

However, the performance of natural action sequences, which is referred to as “script execution” in this paper, is frequently disrupted by unpredicted events or distractions unrelated to the goal. Little is known about preschoolers’ abilities to control script execution endogenously in such non-routine situations. Yet, some recent studies have tackled this issue. For example, when observing a misleading demonstration of familiar action sequences that included irrelevant actions, 5-year-olds avoided overimitating the irrelevant actions, but 3-year-olds tended not to (Freier et al., 2015). Five-year-olds can also modify familiar scripts when necessary. Yanaoka

(2014) developed a “doll task,” which measures the ability to execute a script for changing clothes. When presented with a doll wearing clothing items incorrectly, 5-year-olds could remove them until the familiar script can be followed. Thus, young children not only acquire event knowledge but they can also gradually acquire the ability to control script execution; however the control processes underlying script executions in such non-routine situations have yet to be fully understood.

To date, several previous studies have demonstrated many models of controlling script execution in non-routine situations (e.g., Cooper, Ruh, & Mareschal, 2014; Cooper & Shallice, 2006; Trafton, Altmann, & Ratwani, 2011; Wood & Neal, 2007). In these models, the goal concept is central to the control of script execution. For example, Cooper et al. (2014) recently developed a “goal circuit model,” which is based on the Dual Systems framework (Norman & Shallice, 1986). In the model, the activation of goal units supports the function of the supervisory system¹, which helps us to add top-down control for cases in which action is not fully routinized. Consistent with this theory, most of studies regarding executive functions have focused on goal maintenance and demonstrated its role in top-down control (e.g., Miyake & Friedman, 2012; Munakata, Herd, Chatham, Depue, Banich, & O’Reilly, 2011). Moreover, it is important to note that scripts consist of both main goals and sub-goals set up hierarchically. Indeed, script execution is controlled at the multiple goal level (Cooper et al., 2014). Based on this theoretical framework (e.g., Cooper et al., 2014), it has been hypothesized that in controlling script execution in non-routine situations, young children maintain their hierarchical goal representations, detect the conflict with their goals, and resolve it. Thus, the improvement in the ability to maintain a hierarchical goal representation is essential for the development of script

¹ The supervisory system is like an executive system in that its function consists of several elements, such as goal generation, error monitoring, and strategy generation (e.g., Shallice & Cooper, 2011).

execution control. Therefore, in this study, we explored preschoolers' ability to maintain a hierarchical goal representation in the control of script execution with particular emphasis on its relationship with executive functions. In the following sections, we discuss the theoretical significance of "maintenance of hierarchical goal representations" and "executive functions" to explore the control process underlying script execution in young children.

A number of studies have indicated that scripts are hierarchically represented from a main goal to sub-goals (e.g., Lashley, 1951; Schank & Abelson, 1977; Zacks & Tversky, 2001). A simulation model of script execution was proposed, assuming the goal hierarchy of action sequences (Cooper et al., 2014). Executing scripts based on sub-goals is very effective to complete a fixed action sequence because sub-goals are hierarchically structured to attain a main goal (Cooper et al., 2014). For example, we can attain a main goal (making coffee) by continuously performing sub-goals in a fixed order (adding coffee grounds, sugar, and cream). Each sub-goal constrains subsequent actions and enables us to detect goal-irrelevant information and resolve the situation (Cooper et al., 2006). However, as predicted from Cooper et al. (2014), maintaining a sub-goal causes us to rigidly execute action sequences because it inhibits the flexible selection of other optional sub-goals. Here, the function of a main goal involves activating and coordinating appropriate sub-goals. Script execution based on a main goal enables us to flexibly select optional sub-goals for attaining the final goal (Cooper et al., 2014). For example, in maintaining a main goal (making coffee), we can select two optimal sub-goals such as adding sugar from a pack and adding cream, in variable order. Munakata, Snyder, and Chatham (2012) also suggested that young children gradually acquired the ability to maintain abstract goal representations (going to the grocery store), which support the flexible selection of sub-goals (buying milk, bread, and chocolate). These main goal/sub-goal hierarchies allow us to

perform a variety of action sequences and adapt to diverse environments (e.g., Ruh, Cooper, & Mareschal, 2010).

The management of a hierarchical goal representation is very demanding particularly for younger children (e.g., Amso, Haas, McShane, & Badre, 2014). For example, three- and four-year-old children have difficulty maintaining a main goal (Chevalier & Blaye, 2008; Marcovitch, Boseovski, Knapp, & Kane, 2010; Towse, Lewis, & Knowles, 2007). However, 5-year-olds progressively develop the ability to maintain a main goal (e.g., using three colors equally for coloring six animal shapes) and endogenously adjust action sequences in a relatively long sequential task (Freier, Cooper, & Mareschal, in press; Shapiro & Hudson, 2004). Moreover, Freier et al. (2015) showed that 5-year-olds successfully controlled their imitation of familiar action sequences to avoid irrelevant actions that had been inserted. Thus, we can assume that 5-year-olds might control script execution by maintaining a main goal. However, no study has directly explored the link between the control of script execution and the ability to maintain hierarchical goal representations, partly because it is difficult to find an appropriate natural sequential task which is familiar to children.

Therefore, our first aim was to examine through a novel paradigm, whether young children could control script execution by maintaining hierarchical goal representations. For this, we modified a “doll task” (Yanaoka, 2014) in which each child is instructed to help a doll put on seven items (shoes, shirt, blazer, trouser, socks, underpants, and school bag) to attend kindergarten. This task was devised to measure the preschoolers’ abilities to control script execution. Specifically, we focused on the order of the items involved in changing clothes. In the new doll task, we prepared two types of order relations: “invariant order” and “variant order” (Appendix 1). For example, putting on a blazer and then a school bag represents an “invariant

order” relation because the donning of a blazer always precedes the donning of a school bag. In contrast, putting on a blazer and then trousers represents a “variant order” relation because it is possible to put on either a blazer or trousers after putting on a shirt and underpants. In other words, invariant order items are always unchangeable, whereas the donning order of the items in the variant order category is exchangeable. Previous studies have shown that the orders of successive actions within a script are not all equal in strength (Botvinick & Plaut, 2004; Price & Goodman, 1990; Ratner, Smith, & Dion, 1986; Ruh et al, 2010). Thus, this type of order manipulation is quite natural for a script occurring in children’s daily lives.

In this new doll task (hereafter the *closet doll task*), 4-and 5-year-olds were asked to judge whether they should put on the items from each shelf of the closet one by one. Then they encountered order errors of the items in the invariant order category (e.g., trousers appearing before underpants). In order to dress the doll in the appropriate order for attending kindergarten, they were required to correct the wrong order of the invariant order items. That is, they had to decide not to use an item shown in the wrong order (to skip the item) and to move to the next shelf for the appropriate item. This skip-and-move behavior in the face of the invariant order errors is compatible with both the main goal (changing clothes to attend kindergarten) and the sub-goals (putting on the next appropriate items). It is also important to examine the items in the variant order category. By definition, the order of items in the variant order category is changeable. However, children might expect the next item to be on the next shelf. This expectation could be a sub-goal of the next action. If the children encounter an unexpected item that conflicts with the sub-goal, they might unnecessarily skip the item and move to the next shelf even when the item is from the variant order category. In contrast, if the main goal is dominant in the maintenance of a hierarchical goal representation over the sub-goal, they might

avoid skipping the items unnecessarily, and accept the unexpected order of the variant order items to dress the doll. Thus, it is critical to examine children's skip-and-move responses to items from the variant order category in a situation wherein the main goal and a sub-goal can potentially lead to different actions. Finally, younger children who could not maintain either the main goal or sub-goal may accept any order of variant items and invariant order errors.

In order to directly assess the status of a hierarchical goal representation in the script execution control, we manipulated the strength of sub-goal maintenance with the presence or absence of visual and verbal labels; i.e., asking children what they will do next. The presence of the labels creates a sub-goal dominant situation in the maintenance of the hierarchical goal representation. In the closet doll task, the sub-goal maintenance supported by the labels would prompt children to expect the next appropriate items. In the label group, children would then encounter an inconsistency between the label and the presented items in both the invariant and variant order category. Hence, the label manipulation should cause children to skip the target items in both invariant and variant order category. We predicted that the dominant status of the sub-goal maintenance would be achieved by the presence of the labels irrespective of the prior / original status of the hierarchical goal representation (i.e., main goal dominant, sub-goal dominant, or weak goal activation).

Our second aim was to determine which factors of executive functions explain abilities to control script execution beyond age differences. It has repeatedly been demonstrated that executive functions—the processes involved in the conscious control of thought and goal-directed behavior—develop dramatically during the preschool years (e.g., Garon, Bryson, & Smith, 2008). Executive functions are conceptualized as consisting of dissociable factors: updating (i.e., holding and manipulating information in working memory), inhibition (i.e.,

suppressing prepotent responses), and shifting (i.e., switching flexibly between tasks or mental sets) (e.g., Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000). Although these three factors have different roles in explaining individual differences in executive function tasks, they involve a shared unitary construct (common executive function); one candidate for the shared construct among the three factors is the ability to maintain goals (e.g., Friedman, Miyake, Young, DeFries, Corley, & Hewitt, 2008; Miyake & Friedman, 2012). Moreover, regarding the relationships between the common executive function and the three factors (inhibition, updating, and shifting) proposed by Miyake et al (2000), once the correlations with the common executive function factor were controlled, shifting and updating tasks captured updating-specific and shifting-specific factors respectively; whereas inhibition tasks correlated with the common executive function virtually perfectly (Friedman et al., 2008). They suggested that a factor behind inhibition might reflect goal maintenance more clearly than shifting and updating (Miyake & Friedman, 2012; Munakata et al., 2011). For example, Towse et al. (2007) indicated that goal neglect, the failure to represent the mapping between cues and actions with sufficient saliency in appropriate timing (Kane, Conway, Hambrick, & Engle, 2007), was uniquely linked to the inhibition abilities of preschool children. Moreover, Barker and Munakata (2015) demonstrated that prompting 3-year-olds to activate and maintain goal representations through reminders facilitated their inhibitory control. Thus, we hypothesized that if the endogenous control of script execution required the maintenance of either sub-goals or main goals, inhibition ability would be more strongly related to the control of script execution than updating or shifting. Moreover, this hypothesis led to another prediction: Prompted sub-goal maintenance through labels—which could reduce the load of sub-goal maintenance—should reduce the influence of executive functions on the endogenous control of script execution.

In this study, using the novel paradigm, we aimed to examine the preschoolers' abilities to control their script execution based on the maintenance of hierarchical goal representations. Table 1 describes the overview of our experiments with predictions. In two experiments we manipulated the strength of sub-goal maintenance through the use of visual labels (Experiment 1) and verbal labels (Experiment 2). Moreover, we explored the association of abilities to control script execution with well-established executive function tasks. This multi-task approach helps us to not only specify the control process involved but also to validate the abilities measured by the novel paradigm.

 Insert Table 1 about here

EXPERIMENT 1

Methods

Design

We employed a 2 (Age Group; 4-year-old and 5-year-old) \times 2 (Condition; label and control) experimental design. In order to capture the developmental changes in the abilities to maintain the sub-goal and/or the main goal, we tested 4- to 5-year-olds. It is known that most 3-year-old children are susceptible to the effects of goal-irrelevant information (Freier et al., 2015); thus, they were considered unsuitable for the present study. In the label group, we provided a visual label (a picture) on each shelf of the closet, expecting to support the maintenance of sub-goals. In the control group, no labels were provided. The children were randomly assigned to either the label group or the control group within each age group.

Participants

Eighty-eight children (46 boys and 42 girls) attending two kindergarten schools in the cities of Kyoto and Wakayama participated. Two children could not participate during the second data collection period. Forty-four 4-year-old children ($M = 51.78$ months, $SD = 3.56$ months, $range = 46\text{--}59$ months) and forty-two 5-year-old children ($M = 67.01$ months, $SD = 3.80$ months, $range = 60\text{--}73$ months) were included in the subsequent analyses.

All children who took part in the two sessions had experienced “changing clothes” to attend kindergarten for more than one year; thus, they were familiar with the “changing clothes” script. No children were reported as having developmental atypicalness. Their socioeconomic background was predominantly middle class. Informed consent was obtained from the parents or the kindergarten staff members on behalf of all the children prior to their participation in the study. The study was approved by the institutional ethics committee of the Graduate School of Education, Kyoto University.

Procedure

The individualized experiment was divided into two sessions. To put the children at ease, both sessions were conducted in a quiet room at the kindergarten. In the first session, we conducted four tests to assess their vocabulary comprehension and executive functions, as follows: the Japanese version of the Picture Vocabulary Test (Ueno, Nagoshi, & Konuki, 2008), the red/blue task (revised Simpson & Riggs, 2005), the Dimensional Change Card Sort task (Zelazo, 2006), and the Nine Box task (Wiebe, Sheffield, Nelson, Clark, Chevalier, & Espy, 2011). The testing required approximately twenty-five minutes to complete. In the second

session, which was conducted 1 week after the first session, the closet doll task was used to measure the ability to control script execution. This task required approximately 15 minutes.

Measures

Picture Vocabulary Test (Ueno et al., 2008)

This test was designed for children who ranged in age from 3 years to 12 years and 3 months; it was developed to assess vocabulary comprehension. It is a multiple-choice test in which a child selects one picture that corresponds to an orally presented word among four pictures on each page. The scope of the testing is determined by the child's capabilities, and its maximum score is 89. If a child had fewer than 3 correct answers on one page and fewer than 2 correct answers on the next page, the test was terminated. This test was performed to reveal the language-specific effects of executive function on the ability to control script execution, as it is known that executive functions are closely related to language development (Cragg & Nation, 2010).

Executive Functions

Red/blue task. This task was conducted to measure inhibition abilities. The procedure was essentially the same as the black/white task used by Simpson and Riggs (2005). During the warm-up phase, the children were presented with both a red card and a blue card, and they were asked to point to the card whose color the experimenter orally indicated. All the participants were able to point to the correct cards. The experimenter subsequently provided the following task instruction: "We are going to play a 'silly game!' In this game, when I say 'red,' I want you to point to the 'blue' card, and when I say 'blue,' I want you to point to the 'red' card." In the pre-test phase, the experimenter told the participants either 'red' or 'blue,' and they were required to point to the opposite-color cards. If they answered correctly, the experimenter praised them and

continued to engage in the practice trial with the other color. If they answered incorrectly, the experimenter explained the task instruction again. This procedure was repeated until each child answered two practice trials in succession. The test phase consisted of 14 trials, during which the experimenter said ‘red’ and ‘blue’ seven times each in a pseudo-random sequence. The total score consisted of the number of correct trials during a session (0–14).

DCCS. This task was conducted to measure shifting abilities. The procedure closely followed the protocol described by Zelazo (2006). The children were presented with two trays and two target cards. The experimenter labeled the target cards using both dimensions, as follows: “Here is a yellow cup, and here is a green ship.” After the labeling, the experimenter provided the task instruction: “We are going to play a card game. This is the color game. In the color game, you should sort the cards according to their colors. If it is yellow, it goes to this tray; and if it is green, it goes to this tray.” The experimenter demonstrated the sorting once and subsequently provided the children the opportunity to sort the next test card with feedback. If they answered correctly, the pre-switch phase immediately followed. For each trial, the experimenter presented a test card and asked the participants to sort the cards according to the relevant rule. Six test cards were presented in a pseudo-random sequence, and no feedback was provided. Once the pre-switch phase was completed, the experimenter introduced the children to a new game in which they were instructed to sort the test cards according to the other dimension (i.e., the dimension switch instruction was provided). In the post-switch phase, the children were presented with six test cards and were asked to sort them according to the new rule. The total score represented the number of correct trials only during the post-switch phase.

Nine Box task. This task was conducted to measure updating abilities and was adapted from Diamond, Prevot, Callender, and Druin (1997). The procedure for this task was very similar

to that used by Wiebe et al. (2011). At the start of the trials, the children were presented with nine boxes with different patterns and colors. The experimenter placed a marble in each box, and then closed all nine lids. The boxes were then pushed forward toward the children, and they were required to find the marbles hidden in all the boxes. They were allowed to reach to any box in any order, but they could open only one box during a trial. They were also instructed to remember which box contained a marble without receiving any information regarding the placement of the boxes. After they opened the box, the children received feedback (“You can find it,” or “There is nothing in there”), and the box was then recovered. A delay period of five seconds began from the moment the opened box was recovered. During the delay period, the children were asked to close their eyes for five seconds, while the experimenter scrambled all the boxes in a random order. This procedure was repeated for 20 trials, or until they found all the marbles. The score was measured as the longest run of consecutive correct responses.

Closet Doll Task

Materials and procedure. This task was adapted from the doll task (Yanaoka, 2014). As shown in Figure 1, a closet with seven shelves was prepared for storing the seven items. Hence, in this task, a main goal corresponded to dressing to attend kindergarten, and a sub-goal corresponded to donning the next appropriate item. This task consisted of one script-confirmation block and three item-change blocks.

Insert Figure 1 about here

The young children initially engaged in the script-confirmation block to enable an assessment of their abilities to execute scripts in routine situations. They were shown the closet

with the visual labels and the doll, whose name was Naomi for the girls and Ken for the boys. The visual labels depicted the items necessary to attend kindergarten, and they were attached to the shelves. The experimenter introduced the children to the names of all the items twice and provided the following instructions: “Naomi (Ken) is very careful. She (He) dons seven items in the closet carefully, and on each shelf she (he) attaches the visual labels allowing her (him) to know which items are on each shelf. I want you to help Naomi (Ken) choose the items and dress to attend kindergarten.” After the children were asked to decide which items they would put on next and provided their answers, the experimenter opened the shelf and encouraged them to dress the doll. This procedure was repeated until they had finished dressing the doll.

In the item-change block, the children received practice trials in which they were required to decide whether they could put shoes, a school bag, and underpants on a naked doll, and we confirmed that all children judged the three items appropriately. After the practice trials, we proceeded to the test trials, in which another doll whose name was Sarah (or Joe, for the boys) and who had already worn both underpants and a shirt (for List A, see Table 2) or a shirt (List B) or underpants (List C) was introduced. The experimenter subsequently opened each shelf individually, and the children were asked whether they would put the item on the doll for the purpose of attending kindergarten. If they decided not to put the item on, the experimenter returned it to the shelf. This procedure was repeated until they had dressed the doll completely. Three lists of the items (see Table 2) were used in the counterbalanced order that included both an “invariant order” and a “variant order.” The invariant order implied that the dressing order of the items was not interchangeable; for example, wearing underpants on trousers was incorrect, based on the changing clothes script. In contrast, the variant order implied that the items were

interchangeable; for example, we could choose shirts, socks, or trousers after we put on the underpants.

The difference between the control and label groups was whether the experimenter asked the children which items they would or would not select among the visual labels. In the control group, the children were provided with the following instruction: “Sarah is very careless. She forgets to attach these visual labels to each shelf, so we do not know what items are in the shelves. I want you to help Sarah choose the items and put them on so she can attend kindergarten.” No visual labels were attached to any shelf, although the children were allowed to see the visual labels spontaneously left on the desk to confirm what items were available. The experimenter opened each shelf from the top, and the children were asked whether they would like to dress the doll with that item. In the label group, the children were provided with the following instruction: “Sarah is very careless. She often mistakenly attaches these visual labels to each shelf, so items in the shelves might be different from the visual labels. I want you to help Sarah choose the items and put them on so she can attend kindergarten.” The visual labels were attached to each shelf. The children were asked to select which item to put on their doll by pointing to the visual label before the experimenter opened the shelf. The other procedures were the same as for the control group. As shown in Table 2, the visual labels were inversely attached to conflict with the invariant and variant orders.

Coding. We employed two types of order to measure performance in the item-change block. The first index was the invariant order, especially focusing on whether children could skip target items in the invariant order category. The target item was an invariant order item shown in the wrong order (trousers appearing before underpants). The first index was coded as whether children skipped the target item or not (skip trousers). The maximum score was 3 because each

of three lists included one wrong order with the invariant order items. To attain the main goal (to correctly don the items to wear to attend kindergarten) or the sub-goal (to don the next appropriate item), children needed to overcome the deviation from the invariant order. This first index reflected detecting and resolving the conflict with the main goal or sub-goals.

The second index was the variant order, especially focusing on whether children skipped the target items in the variant order category. Each list included two or three variant order items, and the second index was whether children skipped at least one variant order item in each list. Thus, the maximum score was 3. In the case of the variant order items, the children did not have to stop to put them on in light of the main goal. However, if they assumed the next item following underpants should be a shirt (i.e., they held to a sub-goal), they might delay putting on the other variant items (socks or trousers) until the shirt appeared. This second index reflected detecting and resolving conflict based on the sub-goals rather than the main goal. Both indexes reflected achieving the ability to control script execution based on hierarchical goal representations (see Table 1).

All data were coded based on a video recording by the experimenter and an undergraduate psychology major. For the index of invariant order, the consistency rate was 96% ($k = .92$), and for the index of variant order, the consistency rate was 100% ($k = 1$). The two individuals resolved coding disagreements through discussion.

Insert Table 2 about here

Results and Discussion

Closet Doll Task

Script-confirmation block. This block aimed to determine whether young children could execute the “changing clothes” script without errors. Overall, 6 of the 86 children who took part in two sessions did not pass the script-confirmation task. Two of the children belonged to the five-year-old group, and the other four children were four years old. Consistent with previous studies (Freier et al., 2015; Yanaoka, 2014), the majority of children in this age range were capable of executing “changing clothes” scripts in non-conflict situations. Following Yanaoka (2014), in this study we analyzed 80 young children after excluding the six children, with 20 participants in each condition.

Item-change blocks. Item-change blocks aim to clarify the development of preschoolers’ abilities to control script execution in non-routine situations. The blocks resulted in two indexes measuring how young children controlled two kinds of order reflecting how often they skipped target items in the invariant order category and the variant order category.

Insert Table 3 about here

Table 3 summarizes the means and standard deviations. In terms of the invariant order, an ANOVA using age group (between: four-year-olds or five-year-olds) and condition (between: control or label) was conducted. We identified significant main effects of condition, $F(1, 76) = 14.13, p < .001, \eta_p^2 = .16$, and age group, $F(1, 76) = 13.02, p = .001, \eta_p^2 = .15$, which indicated that children in the label condition skipped target items from the invariant order category more than those in the control condition, and that the five-year-olds skipped target items from the invariant order category more than the four-year-olds. Furthermore, a significant interaction

between the factors was found, $F(1, 76) = 13.02, p = .039, \eta_p^2 = .06$. According to the post-hoc analyses of the age difference in each group, the five-year-olds skipped target items from the invariant order category more than the four-year-olds did in the control condition ($t(76) = 4.12, p < .001, d = 1.04$), but not in the label condition ($t(76) = 1.18, p = .241, d = .13$). Moreover, the post-hoc analyses of the label effects on each age group showed that the visual label facilitated the performance of the four-year-olds ($t(76) = 4.17, p < .001, d = 1.09$), but not of the five-year-olds ($t(76) = 1.04, p = .303, d = .12$).

The same analysis for the variant order items subsequently indicated a significant main effect of condition, $F(1, 76) = 40.77, p < .001, \eta_p^2 = .35$; however, there was no significant main effect of age group $F(1, 76) = 13.02, p = .001, \eta_p^2 = .15$, which indicated that children in the label condition skipped target items from the variant order category more than those in the control condition. In addition, there was no significant interaction between the factors, $F(1, 76) = 0.90, p = .765, \eta_p^2 = .001$. These results suggest that for both age groups, rigidity in executing scripts is strengthened by sub-goal maintenance, which is assumed to be facilitated through the use of visual labels.

These findings regarding invariant and variant order indicated developmental changes in the performance of the closet doll task. In our sample, 4-year-olds were insensitive to the deviation from an invariant order; however, the label manipulation that promoted sub-goal maintenance caused the children to skip target items from both invariant and variant order categories. By contrast, as predicted from Cooper et al.'s model (2014), the sub-goal maintenance caused most 5-year-olds to unnecessarily skip target items from the variant order category without apparent influences on performance of invariant order items. Conversely, in the control condition, most 5-year-olds could endogenously modify the irregularity of items in the

invariant order category and accommodate the orders of the variant items; thus, they were able to maintain a main goal and flexibly coordinate optional sub-goals.

Executive Function and Sub-goal maintenance in the Item-change blocks

The results of the vocabulary test and the executive function tasks are shown as a function of age (Appendix 1). We subsequently analyzed the DCCS², and the distribution was not normal.

Children were considered to have passed the second dimension if they correctly sorted five of the six cards (Kirkham, Cruess, & Diamond, 2003).

Our second aim was to demonstrate the extent to which executive functions explained the ability of preschool children to control script execution. We examined whether performance on the item-change blocks was related to performance on executive function tasks after controlling for age and vocabulary tests. As shown in Table 4, a correlation analysis was conducted for each group. Notably, even after controlling for age and vocabulary, the number of skips of the target items in the invariant order category was significantly related to the red/blue task for the control group; however, for the label group, there was no significant relationship between the item-change and executive function tasks. The number of skips of the target items in the variant order category was not related to executive function tasks.

 Insert Table 4 about here

Moreover, to determine the specific contribution of the interaction between condition and the executive function tasks, we conducted a series of hierarchical multiple regressions. In

² We determined that 22 of the 40 four-year-olds were in the correct answer group, and 34 of the 40 five-year-olds were in the correct answer group. We conducted a chi-square test to examine age-related effects and identified a significant main effect ($p < .01$).

the first step, with the number of skips of target items in the invariant order category from the item-change blocks as the dependent variable, we included age and vocabulary task performance. In the second step, three executive function tasks and the condition (control = 0, label = 1) were entered. In the third step, the interactions between the executive function tasks and the condition were entered. The results are summarized in Table 5.

 Insert Table 5 about here

Consistent with the partial correlation analyses, both the condition ($\beta = .53, t = 3.17, p = .002$) and the interaction between the condition and the red/blue task ($\beta = -.31, t = -2.65, p = .010$) were significant predictors of the number of skips in the invariant order category. According to the post-hoc analysis, in the control group, the red/blue task was a significant factor ($\beta = .31, t = 3.32, p = .001$) but it was not in the label group ($\beta = -.19, t = -0.26, p = .798$). The red/blue task was closely related to the number of skips in the invariant order category—both required goal maintenance (i.e., the task goal for the red/blue task and sub-goal for skipping the wrong items in the invariant order category). The use of the visual label, which is assumed to promote sub-goal maintenance, may help children detect and resolve conflict during script execution, removing the inter-individual variations in the number of skips of items from the invariant order category and consequently wiping out its correlation with red/blue task performance.

The same analyses were conducted for the variant order items. The overall regression model was also significant, $F(9, 70) = 5.61, p < .001$, and the independent variables explained 42.3% of the variation in performance of the item-change blocks. Only the condition factor ($\beta =$

1.32, $t = 5.76$, $p < .001$) was a significant predictor of the number of skips of the target items in the variant order category, indicating that sub-goal maintenance strengthened the rigidity in script execution, regardless of executive function development.

Developmental Changes in the Control of Script Execution Based on Sub-goals or a Main Goal

Although our results suggest that 5-year-olds tend to control script execution based on a main goal and most 4-year-olds cannot control it, we also observationally noted that some children endogenously maintained the sub-goals, indicated by the fact that they skipped items from both the invariant order category and variant order category, even in the control group. By using the number of skips of items in the variant order category as a measure, we analyzed whether children controlled script execution based on sub-goals or a main goal. We confined this analysis to children ($N = 66$) who succeeded in least two of the three trials requiring skips of items in the invariant order category, because they could correct the invariant order errors based on a main goal or sub-goals. When we tested the correlation between age and the number of skips of items in the variant order category in each condition, we found, intriguingly, that in the control group there was a significant negative relation between age and the number of skips ($r(27) = -.47$, $p = .014$). However, there was no relation in the label group ($r(39) = .12$, $p = .482$). These findings indicate that after excluding children who could not skip target items in the invariant order category, younger children are likely to skip target items in the variant order category; whereas older children tend to accommodate variant order.

We also examined the correlation between the number of skips of items in the variant order category and executive functions, controlling for age and vocabulary, and found that inhibition was significantly related only in the control condition after excluding children who could not skip target items in the invariant order category ($r(23) = -.45$, $p = .023$). Although the

sample was limited, there was a possibility that inhibition was related to the ability to stop setting only one sub-goal endogenously.

Limitations of Experiment 1

Experiment 1 has two limitations in its design. First, the visual labels are so salient that children may strongly face conflicts with the different items on the shelves, thus leading them to skip responses. Indeed, Shapiro and Hudson (2004) reported that displaying a picture of a finished product enabled children to execute scripts without making mistakes. In Experiment 2, we investigated the role of sub-goal maintenance by asking children to verbalize which item they would use to dress the doll, thus eliminating the influence of visual labels. Second, our item list in Experiment 1 included three trials; however, one list (List B) was not rigidly counterbalanced regarding the variant items. In Experiment 2, as a result of the limit of time taken up by the experiment, we only performed two trials that were perfectly counterbalanced.

EXPERIMENT 2

The main purpose of Experiment 2 was to examine preschoolers' abilities to control script execution based on hierarchical goal representations through the manipulation of verbalized labels. It also sought to replicate the positive association between executive functions, particularly inhibition, and the control of script execution.

Methods

Design

Experiment 2 had the same 2 (Age Group; 4-year-old and 5-year-old) \times 2 (Condition; label and control) design as Experiment 1 but did not employ visual labels.

Participants

A total of 86 children (44 boys and 42 girls) attending two kindergarten schools in the cities of Kyoto and Wakayama were assigned to either the control or label conditions. Forty-three 4-year-old children ($M = 52.65$ months, $SD = 3.75$ months, $range = 46\text{--}59$ months) and forty-two 5-year-old children ($M = 64.95$ months, $SD = 4.19$ months, $range = 60\text{--}73$ months) were included in the final analyses. One child had difficulty in continuing with this experiment.

As in Experiment 1, all children were familiar with the “changing clothes” script. No child was reported to be developmentally atypical. They predominantly had a middle class socioeconomic background. Informed consent was obtained from the parents prior to their children’s participation in this study. This experiment was approved by the same ethics committee at Kyoto University that approved the first one.

Procedure

We proceeded with the same tasks as in the first experiment, except that we did not administer the vocabulary test because it did not predict script execution performance after controlling for the children’s ages. The testing was again conducted in a quiet room in the kindergarten and required approximately thirty minutes for completion.

Measures

Executive functions. We conducted the same three tasks to measure executive functions as in Experiment 1.

Closet doll task. In Experiment 2, the materials and procedures were nearly the same as in Experiment 1, with two exceptions. First, the method of sub-goal maintenance manipulation was based on verbal labels. In control and label groups, in Experiment 2 no visual labels were attached to the shelves and the experimenter opened the shelves from the top. In the label group, children were required to verbalize which items they would select before the experimenter

opened each shelf, instead of pointing to the visualized labels. Second, to maintain a counterbalanced variant order, we excluded List B used in Experiment 1.

All data were coded via a video recording by the experimenter and the same undergraduate who were involved in Experiment 1, with a consistency rate of 100% ($k = 1$).

Results and Discussion

Closet Doll Task

Script-confirmation block. Five of the 85 children who finished all the tasks did not pass the script-confirmation phase of the task. Two children belonged to the five-year-old group, and the remaining children were in the four-year-old group. As in Experiment 1, we analyzed 80 young children after excluding the five above-mentioned children.

Item-change blocks. We initially examined the preschoolers' abilities to control script execution based on hierarchical goal representations (Table 6). In terms of the invariant order category, an ANOVA using age group and condition identified significant main effects of condition, $F(1, 76) = 14.06, p < .001, \eta_p^2 = .16$, and age group, $F(1, 76) = 7.43, p = .008, \eta_p^2 = .09$, which indicated that children in the label condition skipped target items from the invariant order category more than those in the control condition, and that the five-year-olds skipped target items from the invariant order category more than the four-year-olds. Furthermore, a trend of interactions between the factors was detected, $F(1, 76) = 2.91, p = .092, \eta_p^2 = .04$. The post-hoc analyses of the age difference in each group indicated that the five-year-olds skipped target items from the invariant order category more than the four-year-olds did in the control condition ($t(76) = 4.12, p < .001, d = 1.04$), but not in the label condition ($t(76) = 1.44, p = .152, d = .16$). Moreover, the post-hoc analyses of the label effects on each age group showed that the verbal label facilitated the performance of the four-year-olds ($t(76) = 3.14, p < .001, d = 0.90$), but not

that of five-year-olds ($t(76) = 0.72, p = .472, d = .08$). These findings were consistent with Experiment 1.

The same analysis for the variant order items indicated a significant main effect of condition, $F(1, 76) = 23.39, p < .001, \eta_p^2 = .55$; however, there was no significant main effect of age group, $F(1, 76) = 0.37, p = .547, \eta_p^2 = .01$. In addition, there was no significant interaction between the factors, $F(1, 76) = 0.37, p = .547, \eta_p^2 = .01$. These findings were also consistent with Experiment 1.

Consistent with Cooper et al. (2014), even though verbal labels were used instead of visual labels, we replicated the different effects of sub-goal maintenance, depending on the age groups. As reported by Shapiro and Hudson (2004), however, verbalized sub-goal maintenance was less effective than the visualized sub-goal maintenance in Experiment 1, as indicated by the non-significant (marginal) interaction between age and condition in Experiment 2.

 Insert Table 6 about here

Executive Functions and the Sub-goal maintenance on the Item-change Blocks

The results from the executive function tasks are shown in Appendix 1 as a function of age. Regarding the DCCS, as with the analysis in Experiment 1, we conducted a chi-square test to examine the age-related effect and found a significant main effect ($p < .05$)³.

Second, we aimed to determine which factors of executive functions explained preschoolers' abilities to control script execution. As shown in Table 7, a partial correlation analysis that controlled for age indicated that in the control group, the number of skips of the

³ Here, we found that 21 of the 40 four-year-olds were in the correct answer group, while 34 of the 40 five-year-olds were in the correct answer group.

target items in the invariant order category was significantly associated with the red/blue task; however, in the label group, there was no significant association between the item-change and executive function tasks. In contrast with Experiment 1, there was a significant correlation between the number of skips of the target items in the invariant order category and those in the variant order category only in the label group. One of the reasons for this inconsistency is that some older children in Experiment 1 stopped skipping items in the variant order category in spite of the manipulation of the visual labels; whereas a few children in Experiment 2 did not tend to do so. These results suggest that some children have the ability to maintain a main goal against exogenous sub-goal maintenance by visual label, but they cannot maintain it against sub-goal maintenance by verbal label. Further studies are required to test this possibility.

 Insert Table 7 about here

To determine the specific contribution of the interaction between the condition and executive function tasks to the item-change blocks when controlling for the factor of age, we conducted the same hierarchical multiple regression analysis as was conducted in Experiment 1. The results are summarized in Table 8.

 Insert Table 8 about here

Both the condition ($\beta = .41, t = 3.00, p = .004$) and the red/blue task ($\beta = .12, t = 2.29, p = .025$) were associated with the number of skips of the target items in the invariant order category. In contrast to Experiment 1, the interactions between executive functions and condition

were not significant. These findings suggest that the red/blue task was more closely related to the control of script execution. However, verbalized sub-goal maintenance was less effective than visualized sub-goal maintenance, because the maintenance of verbalized sub-goal maintenance was more demanding. Indeed, the findings of Klingner, Tversky, and Hanrahan (2011), which demonstrated that individuals displayed superior memory recall when items were presented with a picture along with the verbal label compared to when they were presented with only a verbal label, support the results of this study.

The same analyses were conducted for the variant order items. The overall regression model was also significant, $F(9, 70) = 2.84, p = .007$, and the independent variables explained 26.7% of the variation in the item-change blocks. Consistent with Experiment 1, only the condition factor ($\beta = .80, t = 4.61, p < .001$) was associated with the number of skips of the target items in the variant order category, which indicates that sub-goal maintenance strengthens rigidity in the control of script execution.

Developmental Changes in the Control of Script Execution Based on Sub-goal or Main goal

The objective of this analysis was to examine whether children who could skip the target items in the invariant order category controlled script execution at sub-goals level or a main goal level. Therefore, we confined the analysis to children ($N = 67$) who succeeded in more than one of two trials requiring the correction of invariant order errors. We replicated the two findings as in Experiment 1. There was a significantly negative correlation between age and the number of skips of the target items in the variant order category in the control group ($r(28) = -.39, p = .013$), whereas there was no relation in the label group ($r(39) = .11, p = .386$). These patterns from both Experiments 1 and 2 suggest that the younger children in the sample may endogenously set the sub-goal and rigidly detect and resolve conflict during script execution.

Moreover, we considered the correlation between the number of skips of the target items in the variant order category and executive functions. Controlling for the factor of age, there was no correlation with executive function tasks for either condition. In contrast to Experiment 1, inhibition does not explain the ability to stop setting only one sub-goal endogenously.

General Discussion

Previous script studies (Freier et al., 2015; Yanaoka, 2014) have demonstrated that 5-year-olds are able to control script execution in non-routine situations caused by unpredictable events. However, the control processes underlying executing scripts remain unclear. In our study, we highlighted the *maintenance of hierarchical goal representations* in the control of script execution and its relationship with *executive functions*. To understand the function of maintaining hierarchical goal representations in the control of script execution, we manipulated the strength of sub-goal maintenance by helping children select subsequent actions using visual labels (Experiment 1) and verbal labels (Experiment 2).

Development of Hierarchical Goal Representations in the Control of Script Execution

As scripts have hierarchical structures including main goals and sub-goals (e.g., Zacks & Tversky, 2001), young children control script execution based on each goal. According to Cooper et al. (2014), a sub-goal is maintained over a relatively short-term, as it is required moment by moment, but it causes us to rigidly execute a fixed action sequence without order variety. In contrast, the maintenance of a main goal is challenging, as this must occur during the whole script execution process. However, it helps us to flexibly coordinate sub-goals, thus resulting in the successful performance of many different action sequences. During the preschool period, the ability to maintain a goal representation improves (e.g., Blaye & Chevalier, 2011); thus, whether

children can control script execution by maintaining hierarchical goal representations is a fundamental question in developmental psychology.

The two experiments reported here demonstrated that in non-routine situation like the item change blocks, 4-year-olds did not actively maintain a main goal to control script execution; however, they could detect and resolve the order errors of invariant items if they were prompted to maintain a particular sub-goal. This finding is consistent with prior studies demonstrating an inability to maintain an abstract goal in this age group (e.g., Chevalier & Blaye, 2008; Marcovitch et al., 2010). By contrast, five-year-olds were able to endogenously control script execution based on the dominant status of a main goal in the maintenance of a hierarchical goal representation. The fact that they modified the order of invariant items and accommodated the variant order items supports this conclusion. The stronger sub-goal maintenance driven by the use of labels decreased their variant order accommodation performance. These findings contribute to the assessment of the ability to maintain hierarchical goal representations in the control of script execution for each developmental stage. Furthermore, the finding that the maintenance of a main goal helps 5-year-olds flexibly select from appropriate sub-goals is in line with previous studies (e.g., Ruh et al., 2010). These findings are also consistent with the goal circuit model (Cooper et al., 2014), which asserts the importance of goal maintenance in the control of script execution.

Our experiments also revealed an interesting finding. After we excluded the children who could not skip target items from the invariant order category, age was negatively associated with the number of skips of items from the variant order category. This suggests that younger children may endogenously set sub-goals, rather than maintaining the main goal. Previous studies have demonstrated that 6-year-olds who proactively maintain task-relevant information

reacted more quickly and were thus more prepared in a working memory task, but they were more distracted by a secondary task (counting backwards and tapping hands) during the working memory task (Blackwell, Chatham, Wiseheart, & Munakata, 2014; Blackwell & Munakata, 2014). Consistent with these findings, in the control of script execution endogenously setting only one sub-goal involves both the benefit of detecting the wrong order of invariant order items and the cost of overcorrecting optional variant orders. Despite the small sample size of our control group, we could infer that there were developmental changes in the control group in relation to the sub-goal and main goal levels. These findings offer novel insights into the development of the abilities to maintain hierarchical goal representations in the control of script execution.

Control of Script Execution, Goals, and Executive Functions

No empirical study to date has directly demonstrated how young children control script execution in non-routine situations. Executive functions, which are assumed to be composed of at least three factors (inhibition, shifting, and updating), are often linked to the dynamic regulation of goal-directed behaviors. Hence, we also aimed to examine which factors of executive function were associated with the control of script execution in order to clarify its underlying control processes. This investigation is the first attempt to demonstrate that inhibition is associated with the control of script execution in both experiments. A substantial number of studies have indicated that actively maintaining goal representation plays an important role in inhibitory control (Chatham et al., 2012; Munakata et al., 2011; Winter & Sheridan, 2014). Our findings suggest that goal maintenance in the control of script execution is closely related to inhibition. Indeed, in Experiment 1 the sub-goal maintenance boosted by visual labels diminished the positive relationship between inhibition task performance and skips of invariant

items. On the other hand, the impact of the verbal labels in Experiment 2 was not as substantial as the visualized labels, and this was likely because the memory for visualized labels is stronger than that for verbalized labels (Klinger et al., 2011).

However, one question that remains unsolved is why only the factor of inhibition was significantly related to the control of script execution. As Miyake and Friedman (2012) suggested, goal maintenance and implementation are general requirements of all executive function tasks, and common executive function is more purely measured by inhibition tasks than shifting and updating tasks. Towse et al. (2007) identified relationships between “goal neglect” and the red/blue task used in our studies, in which the young children were required to avoid strong prepotent responses. Hence, the effects of updating and shifting involved in the control of script execution may be attenuated when the effect of inhibition is controlled.

Our findings also provide clues that clarify the processes underlying the control of script execution. It seems that sub-goal maintenance helps children detect conflict between the goal and ongoing activities, and that they resolve the conflict during script execution. As Freier et al. (in press) suggested, conflict monitoring theory based on cognitive neuroscience evidence (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001) also supports these assumptions. According to this theory, when conflict in information processing is detected, and a specific brain area (i.e., the anterior cingulate cortex) plays an important role in signaling the conflict, it leads to efficient utilization of cognitive control to resolve the conflict. Furthermore, Decender, van Opstal, and van den Bussche (2014) have shown that the subjective experience of conflict plays a more crucial role than actual response conflict itself in triggering cognitive control. Consistent with this evidence, recent developmental work (Doebel & Zelazo, 2016) has shown that experience with contrastive language, which has the form not X, Y (“No, this is not an apple. It’s different.

It's a banana.”), benefits executive function in 3-year-old children. Thus, the awareness of conflict facilitated by goal maintenance might underlie the control of script execution. These findings and backgrounds are consistent with the theoretical framework provided by the “goal circuit model” (Cooper et al., 2014). In this model, it is assumed that goal representation activates top-down control via the executive system, namely, the supervisory system.

Methodological Implication

Finally, our novel task contributes to research on natural action sequences as it reflects the real lives of preschoolers. The closet doll task consists of a hierarchical structure that young children experience every day. Scripts are composed of a mixture of invariant and variant orders among hierarchically arranged actions (e.g., Price & Goodman, 1990). Previous studies with adult participants (Botvinick & Plaut, 2004; Cooper et al., 2014) have employed a beverage preparation task, which consists of both invariant and variant orders, whereas there are no such appropriate tasks for use in the developmental psychology field. The present study has overcome this limitation by improvising the doll task (Yanaoka, 2014). This ecologically valid task provides opportunities for examining the development of script execution in future research.

Limitations and Conclusion

Despite the notable strengths of our study, it was subject to three limitations. First, we assumed that our closet doll task required children to maintain hierarchical goal representations, but one might argue that they pass the closet doll task by using prior knowledge about clothing instead of maintaining the goal representation. In fact, children were familiar with “clothing script” in the closet doll task. In other words, they might have already acquired the causal knowledge between two pieces of clothing. That knowledge could potentially help children to detect the order errors. Although we cannot dismiss this possibility based on the current data set,

recent papers (e.g., Loucks, Mutschler, & Meltzoff, in press) seem to support our assumption that 5-year-olds have developed the abilities to maintain a main goal, like that available in our task. Loucks et al. reported that even 3-year-olds represented a higher-order goal to organize the memory for others' action sequences, indicating that they can process and represent human action in an abstract fashion. Thus, children older than 3-years tend to represent goal information in executing action sequences. From these pieces of developmental evidence, it is assumed that at least 5-year-olds in this study control script execution based on goal representation rather than the causal knowledge, but future studies are required for further investigation of this issue.

The second limitation involves the measurement of skipping variant order items. Compared with scores on invariant order items, variations in scores on variant order items are higher. It is not possible to identify the cause of the difference in the variations. As we designed two between-participant experiments, some influences of original individual differences might exist. Having both label and control conditions will be necessary in within-participant experiments.

The third limitation involves the measurement of executive functions. We employed only three tasks because of time constraints; however, performance on each of these tasks also reflects non-executive processes, due to the task-impurity problem (Miyake et al., 2000). Therefore, the use of different measures of the same component of executive functions and the extraction of common variance to yield a purer measure of the executive functions is an ideal method (Garon et al., 2008).

Despite these limitations, the current findings clarified the control processes underlying the control of script execution at each developmental stage from the viewpoint of the maintenance of hierarchical goal representations and executive functions. The findings also

offered novel insights regarding how young children control script execution in non-routine situations.

Acknowledgments

The research reported here was supported by a Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) fellows (16J0372). We are grateful to the staff and students of the kindergarten schools for their participation in this research. We thank Yuko Munakata for helpful comments and suggestions on the earlier version of this paper. We also thank four anonymous reviewers for constructive and encouraging comments on the manuscript.

References

- Amso, D., Haas, S., McShane, L., & Badre, D. (2014). Working memory updating and the development of rule-guided behavior. *Cognition*, *133*, 201-210.
- Barker, J. E., & Munakata, Y. (2015). Time isn't of the essence: Activating goals rather than imposing delays improves inhibitory control in children. *Psychological Science*, *26*, 1898-1908.
- Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *53*, 153-164.
- Blackwell, K. A., Chatham, C. H., Wiseheart, M., & Munakata, Y. (2014). A developmental window into trade-offs in executive function: The case of task switching versus response inhibition in 6-year-olds. *Neuropsychologia*, *62*, 356-364.
- Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*, 203-211.
- Blaye, A., & Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *Journal of Experimental Child Psychology*, *108*, 469-483.
- Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, *111*, 395-429.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624-652.
- Chevalier, N., & Blaye, A. (2008). Cognitive flexibility in preschoolers: the role of

representation activation and maintenance. *Developmental Science*, 11, 339-353.

Chatham, C. H., Claus, E. D., Kim, A., Curran, T., Banich, M. T., & Munakata, Y. (2012).

Cognitive control reflects context monitoring, not motoric stopping, in response inhibition.

PLoS ONE, 7, Article e31546.

Cooper, R. P., Ruh, N., & Mareschal, D. (2014). *The goal circuit model: A hierarchical multi-route model of the acquisition and control of routine sequential action in humans*. *Cognitive Science*, 38, 244-274.

Cooper, R.P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, 113, 887-916.

Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics in Cognitive Science*, 2, 631-642.

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science*, 25, 675-683.

Diamond, A., Prevor, M., Callender, G., & Druin, D. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU—introduction. *Monographs of the Society for Research in Child Development*, 62, 1-208.

Doebel, S., & Zelazo, P. D. (2016). Seeing conflict and engaging control: Experience with contrastive language benefits executive function in preschoolers. *Cognition*, 157, 219-226.

Fivush, R. (1984). Learning about school: the development of kindergartners' school scripts. *Child Development*, 55, 1697-1709.

Freier, L., Cooper, R. P., & Mareschal, D. (2015). The planning and execution of natural sequence actions in preschool years. *Cognition*, 144, 58-66.

- Freier, L., Cooper, R. P., & Mareschal, D. (in press). Preschool children's control of action outcomes. *Developmental Science*. doi.org/10.1111/desc.12354
- Friedman N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology General*, 137, 201-225.
- Furman, L. N., & Walden, T. A. (1990). Effects of script knowledge on preschool children's communicative interactions. *Developmental Psychology*, 26, 227-233.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134, 31-60.
- Hudson, J. A., Shapiro, L. R., & Sosa, B. B. (1995). Planning in the real world: Preschool children's scripts and plans for familiar events. *Child Development*, 66, 984-998.
- Hudson, J. A., & Fivush, R. (1991). Planning in the preschool years: The emergence of plans from general event knowledge. *Cognitive Development*, 6, 393-415.
- Hudson, J. A., & Slackman, E. A. (1990). Children's use of scripts in inferential text processing, *Discourse Processes*, 13, 375-385.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in Working Memory*. NY: Oxford.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension switching task. *Developmental Science*, 6, 449-467.
- Klinger, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, 48, 323-332.

- Lashley, K. (1951). The problem of serial order in behavior. In L. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: Wiley.
- Loucks, J., & Meltzoff, A. N. (2013). Goals influence memory and imitation for dynamic human action in 36-month-old children. *Scandinavian Journal of Psychology*, 54, 41-50.
- Loucks, J., Mutschler, J., & Meltzoff, A. N. (in press). Children's representation and imitation of events: How goal organization influences 3-year-old children's memory for action sequences. *Cognitive Science*.
- Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4-to 6-Year-Old Children. *Child Development*, 81, 1687-1695.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Current Directions in Psychological Science*, 21, 8-14.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49-100.
- Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, 15, 453-459.
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: three key transitions. *Current Directions in Psychological Science*, 21, 71-77.
- Nelson, K. (1986). *Event knowledge: Structure and function in development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwarz & D. Shapiro (Eds.), *Consciousness and Self Regulation* (Vol. 4, pp. 1-18). New York, NY: Plenum.
- Price, D., & Goodman, G. (1990). Visiting the wizard: Children's memory for a recurring event. *Child Development*, 61, 664-680.
- Ratner, H. H., Smith, B. S., & Dion, S. A. (1986). Development of memory for events. *Journal of Experimental Child Psychology*, 41, 411-428.
- Ruh, N., Cooper, R. P., & Mareschal, D. (2010). Action selection in complex routinized behaviors. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 955-975.
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum.
- Shallice, T. & Cooper, R. P. (2011). *The Organisation of Mind*. Oxford University Press, Oxford, UK.
- Shapiro, L. R., & Hudson, J. A. (2004). Effects of internal and external support on preschool children's event planning. *Journal of Applied Developmental Psychology*, 25, 49-73.
- Simpson, A., & Riggs, K. J. (2005). Inhibitory and working memory demands of the day-night task in children. *British Journal of Developmental Psychology*, 10, 1-17.
- Towse, J. N., Lewis, C., & Knowles, M. (2007). When knowledge is not enough: the phenomenon of goal neglect in preschool children. *Journal of Experimental Child Psychology*, 96, 320-332.
- Trafton, J. G., Altmann, E. M., & Ratwani, R. J. (2011). A memory for goals model of sequence errors. *Cognitive Systems Research*, 12, 134-143.

- Yanaoka, K. (2014). Effects of planning and executive functions on turning back by young children in the execution of scripts. *Japanese Journal of Developmental psychology*, 25, 232-241. (In Japanese)
- Ueno, K., Nagoshi, N., & Konuki, S. (2008). *Picture Vocabulary Test-Revised*. Tokyo: Nihon Bunka Kagakusha. (In Japanese.)
- Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, 108, 436-452.
- Winter, W., & Sheridan, M. (2014). Previous reward decreases errors of commission on later ‘No-Go’ trials in children 4 to 12 years of age: evidence for a context monitoring account. *Developmental Science*, 17, 797-807.
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, 114, 843-863.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127, 3-21.
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1, 297-301.

Table 1

The overview of the experiments: Manipulation, expected goal status, and predictions

	Condition	Control group		Label group	
	Age group	4-year-olds	5-year-olds	4-year-olds	5-year-olds
Manipulation	Experiment 1	without label		with visual label	
	Experiment 2	without label		with verbal label	
	Expected goal status	Weak	Main goal dominant	Sub-goal dominant	Sub-goal dominant
Prediction	Responses to the invariant order errors	Accept	Skip	Skip	Skip
	Responses to unexpected order of variant items	Accept	Accept	Skip	Skip

Note: The items in invariant order category are always unchangeable for attaining a final goal, but the donning order of the items in the variant order category is exchangeable.

Table 2

Item and Label List in the Closet Doll Task

List A		List B		List C	
Item	Label	Item	Label	Item	Label
Underpants		Shirt		Underpants	
Shirt		Trouser	Underpants	Blazer	Shirt
School bag	Blazer	Underpants	Trouser	Shirt	Blazer
Blazer	School bag	Blazer	Socks	Socks	Trouser
Trouser	Socks	Socks	Blazer	Trouser	Socks
Socks	Trouser	School bag	School bag	School bag	School bag
Shoes	Shoes	Shoes	Shoes	Shoes	Shoes
Label lists were used only in Experiment 1					

Table 3

The Mean Numbers of Skips for the Target Items in the Closet Doll Task in Experiment 1

Condition		Four-year-olds	Five-year-olds
Invariant order	Control group ($N = 40$)	1.50 (1.19)	2.50 (0.67)
	Label group ($N = 40$)	2.53 (0.61)	2.79 (0.42)
Variant order	Control group ($N = 40$)	0.60 (0.82)	0.84 (0.90)
	Label group ($N = 40$)	2.05 (1.09)	2.16 (1.02)

Table 4

Partial Correlation Coefficients Between Scores in the Closet Doll Task and those in Executive-function Tasks in Experiment 1 (Controlled for Age and Vocabulary Test Performance)

	1	2	3	4	5
<i>(a) Control group (N = 40)</i>					
1. Invariant order	-				
2. Variant order	.20	-			
3. Red/blue task	.42**	-.04	-		
4. DCCS	.28	-.03	.48**	-	
5. Nine-box task	-.04	-.19	.40*	.33*	-
<i>(b) Label group (N = 40)</i>					
1. Invariant order	-				
2. Variant order	.03	-			
3. Red/blue task	-.04	-.09	-		
4. DCCS	.17	-.26	.38**	-	
5. Nine-box task	-.22	-.30	.44**	.31*	-

* $p < .05$ ** $p < .01$

Table 5

Effects of Executive Function and Condition on the Closet Doll Task Performance in Experiment 1

Predictor	Invariant order				R^2
	B	95% CI	t		
Step 1					.16**
Intercept	2.11	[1.88, 2.35]	17.98		
Age	0.02	[-.01, .05]	1.55		
Vocabulary	0.01	[-.03, .04]	0.42		
Step 2					.22***
Red/blue task	0.31	[.13, .50]	3.32**		
DCCS	- 0.09	[-.95, .77]	- 0.21		
Nine-box task	- 0.14	[-.29, .01]	- 1.83		
Condition	0.53	[.20, .87]	3.17**		
Step 3					.10**
Red/blue \times Condition	- 0.31	[-.54, -.08]	- 2.65*		
DCCS \times Condition	0.25	[-.76, 1.25]	0.49		
Nine-box \times Condition	0.01	[-.24, .26]	0.11		
Total R^2					.48***

Table 6

The Mean Numbers of Skips for the Target Items in the Closet Doll Task in Experiment 2

Condition		Four-year-olds	Five-year-olds
Invariant order	Control group ($N = 40$)	0.80 (0.89)	1.60 (0.60)
	Label group ($N = 40$)	1.45 (0.51)	1.75 (0.55)
Variant order	Control group ($N = 40$)	0.40 (0.68)	0.40 (0.59)
	Label group ($N = 40$)	1.10 (0.85)	1.30 (0.81)

Table 7

Partial Correlation Coefficients Between Scores in the Closet Doll Task and those in Executive-function Tasks in Experiment 2 (Controlled for Age and Vocabulary Test)

	1	2	3	4	5
<i>(a) Control group (N = 40)</i>					
1. Invariant order	-				
2. Variant order	.21	-			
3. Red/blue task	.43*	.24	-		
4. DCCS	.30	-.18	.38**	-	
5. Nine-box task	.29	.03	.40**	.36*	-
<i>(b) Label group (N = 40)</i>					
1. Invariant order	-				
2. Variant order	.43*	-			
3. Red/blue task	.06	-.09	-		
4. DCCS	-.04	-.03	.40**	-	
5. Nine-box task	-.04	-.04	.56**	.42**	-

* $p < .05$ ** $p < .01$

Table 8

Effects of Executive Function and Condition on the Closet Doll Task Performance in Experiment 2

Predictor	<i>B</i>	95% CI	<i>t</i>	<i>R</i> ²
Step 1				.22***
Intercept	1.21	[1.02, 1.40]	12.66***	
Age	0.02	[-.01, .05]	1.74	
Vocabulary	- 0.01	[-.04, .03]	- 0.17	
Step 2				.15**
Red/blue task	0.12	[.02, .22]	2.29*	
DCCS	0.31	[-.16, .78]	1.30	
Nine-box task	0.11	[-.05, .26]	1.32	
Condition	0.41	[.14, .68]	3.00**	
Step 3				.06
Red/blue × Condition	- 0.05	[-.23, .13]	- 0.53	
DCC × Condition	- 0.35	[-1.07, .38]	- 0.94	
Nine-box × Condition	- 0.16	[.25, -.43]	- 1.15	
Total <i>R</i> ²				.44***

Figure Captions

Figure 1.

Materials used in the closet doll task. The orders of the relationships among the seven items in this task were: ① Shirt→ Blazer→ School bag, ② Underpants→ Trouser→ Shoes, and ③ Socks. Directional markers indicate the causal relationships of order, and the lists of ①, ②, and ③ are in random order.

Figure 1



Appendix 1

Means and SDs for the Vocabulary and Executive-function Tasks (N = 80)

	Four-year-olds (N = 40)	Five-year-olds (N = 40)	<i>F score</i>
<i>Experiment 1</i>			
Vocabulary test	19.65 (4.19)	29.03 (5.61)	72.50***
Red/blue task	9.36 (2.22)	11.53 (1.81)	23.72***
Nine-box task	4.36 (1.38)	5.55 (1.80)	11.75**
<i>Experiment 2</i>			
Red/blue task	10.65 (1.79)	11.75 (1.37)	4.78*
Nine-box task	4.93 (1.23)	5.70 (1.31)	7.49**
* $p < .05$, ** $p < .01$, *** $p < .001$			